EXAMEN FINAL DE PROBABILIDADES Y ESTADÍSTICA I
JUNIO 2014

Realizar las preguntas en hojas separadas, indicando explícitamente todas las fórmulas que se utilicen.
Tanto el alumno que copie como el que se deje copiar no podrá examinarse hasta junio de 2015.

Duración de cada parte: 50 minutos.

Parte 1: UNIDADES DIDÁCTICAS 2 Y 3. Probabilidades con Sucesos y Variables Aleatorias.

1. Un día cualquiera en Nueva York, la probabilidad de que el Hombre de Arena esté haciendo fechorías es del 20 %, la probabilidad de que sea el Doctor Octopus es del 40 % y la probabilidad de que sea el Rino es el resto.

   El número de noticias relacionadas con robos en bancos que se publican durante un mes en el Daily Bugle se distribuye según una Poisson con distinto parámetro para cada villano. Para el Hombre de Arena $\lambda = 3$, para el Doctor Octopus $\lambda = 5$ y para el Rino $\lambda = 4$.

   Peter Parker se da cuenta de que en el último mes se han publicado dos noticias de robos en bancos. ¿Cuál es la probabilidad de que sea el Doctor Octopus quién esté haciendo fechorías?

2. El tiempo que se tarda en contestar las preguntas de un test, en minutos, sigue una distribución normal de media 10 y desviación típica 5.

   a) Sabiendo que nadie ha respondido el test en menos de 2 minutos, ¿cuál es la probabilidad de que un individuo tarde más de 15 minutos en responder al test?

   b) Si no se conociese la información sobre el tiempo mínimo de respuesta, ¿cuál sería dicha probabilidad?

   Recordad que el tiempo es una variable aleatoria que no puede tomar valores negativos.

Parte 2: UNIDADES DIDÁCTICAS 4 Y 5. Inferencia Estadística.

3. El tiempo que se tarda en resolver la primera pregunta de un examen de Probabilidades y Estadística I, en minutos, sigue una distribución normal. Se ha medido el tiempo que han tardado 12 alumnos en resolver dicha pregunta, obteniendo los siguientes valores:

   $$9.3 \quad 8.5 \quad 6.9 \quad 8.8 \quad 7.6 \quad 8.9 \quad 8 \quad 9 \quad 8.1 \quad 9.2 \quad 8.3 \quad 9.2$$

   Obtener un intervalo de confianza para la varianza de dichos datos al 95 % de confianza.

4. Tradicionalmente, una marca de pasta de dientes cuenta con una satisfacción media entre sus clientes de 5.9 en una escala de 7. Se ha cambiado ligeramente el sabor y se ha preguntado a $n = 40$ clientes su satisfacción, obteniendo

   $$\sum_{i=1}^{40} x_i = 253.5 \text{ y } \sum_{i=1}^{40} x_i^2 = 1628.77$$

   Suponiendo que la satisfacción de sus clientes sigue una distribución normal, queremos saber, mediante un contraste de hipótesis, si existen evidencias para considerar que la satisfacción es mayor ahora. Resolverlo con un nivel de significación $\alpha = 0.01$. Calcular también el p-valor e interpretarlo.

SOLUCIONES DEL EXAMEN FINAL
PROBABILIDADES Y ESTADÍSTICA I. JUNIO 2014

Parte 1: UNIDADES DIDÁCTICAS 2 Y 3. Probabilidades con Sucesos y Variables Aleatorias.

1. Definimos los sucesos

$$\begin{aligned} A &= \text{el Hombre de Arena está haciendo fechorías} \\ O &= \text{el Doctor Octopus está haciendo fechorías} \\ R &= \text{Rino está haciendo fechorías} \end{aligned}$$

Por el enunciado se tiene $P(A) = 0.2$, $P(O) = 0.4$ y $P(R) = 0.4$.
Sea, además, X la variable aleatoria que representa el número de noticias publicadas sobre robos en bancos en un mes. Sabemos que

$$X|A \sim \mathcal{P}(3), \ \ X|O \sim \mathcal{P}(5), \ \ X|R \sim \mathcal{P}(4)$$

Nos piden $P(O|X = 2)$, que se obtiene aplicando el teorema de Bayes.

$$P(O|X=2) = \frac{P(O \cap (X=2))}{P(X=2)} = \frac{P(X=2|O)P(O)}{P(X=2|A)P(A) + P(X=2|O)P(O) + P(X=2|R)P(R)}$$

Calculamos las probabilidades que intervienen:

$$\begin{aligned} P(X=2|A) &= \frac{3^2}{2!}e^{-2} = 0.224 \\ P(X=2|O) &= \frac{5^2}{2!}e^{-5} = 0.0842 \\ P(X=2|R) &= \frac{4^2}{2!}e^{-4} = 0.1465 \end{aligned}$$

Sustituyendo, obtenemos

$$P(O|X=2) = 0.24569$$

2. Sea la variable aleatoria $T \equiv$ *tiempo que se tarda en contestar el test (en minutos)*. Se tiene que $T \sim N(10, 5)$.

Como el tiempo debe ser positivo comprobamos si hace falta truncar la v.a. X. En este caso $\mu - 3\sigma = -5 < 0$ con lo que debemos truncar por $T \geq 0$ a la hora de calcular las probabilidades.

a) La probabilidad pedida es

$$\begin{aligned}
P(T > 15|T \geq 2, T \geq 0) &= P((T > 15|T \geq 2) = \frac{P((T > 15) \cap (T \geq 2))}{P(T \geq 2)} = \frac{P(T > 15)}{P(T \geq 2)} = \\
&= \frac{P\left(\frac{T-10}{5} > \frac{15-10}{5}\right)}{P\left(\frac{T-10}{5} > \frac{2-10}{5}\right)} = \frac{P(Z > 1)}{P(Z > -1.6)} = \frac{1 - P(Z \leq 1)}{P(Z \leq 1.6)} = \\
&= \frac{1 - 0.8413}{0.9452} = \frac{0.1587}{0.9452} = 0.1679
\end{aligned}$$

donde $Z \sim N(0,1)$.

b) Si no conocemos la información sobre el tiempo mínimo de respuesta, la probabilidad es

$$\begin{aligned}
P(T > 15|T \geq 0) &= \frac{P((T > 15) \cap (T \geq 0))}{P(T \geq 0)} = \frac{P(T > 15)}{P(T \geq 0)} = \frac{P(Z > 1)}{P\left(Z \geq \frac{0-10}{5}\right)} = \\
&= \frac{0.1587}{P(Z \geq -2)} = \frac{0.1587}{P(Z \leq 2)} = \frac{0.1587}{0.9772} = 0.1624
\end{aligned}$$

Parte 2: UNIDADES DIDÁCTICAS 4 Y 5. Inferencia Estadística.

3. Se pide calcular un IC al 95 % de confianza para la varianza σ^2. La variable pivote que se debe utilizar es:

$$\frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{n-1}$$

Pivotando, se obtiene que el intervalo es:

$$\left(\frac{(n-1)s^2}{\chi^2_{n-1,\frac{\alpha}{2}}}, \quad \frac{(n-1)s^2}{\chi^2_{n-1,1-\frac{\alpha}{2}}} \right)$$

Realizamos algunos cálculos, obteniendo:

$$\sum_{i=1}^{12} x_i = 101.8, \quad \bar{x} = 8.48, \quad \sum_{i=1}^{12} x_i^2 = 869.54$$

$$s^2 = \frac{\sum x_i^2 - n \cdot \bar{x}^2}{n-1} = \frac{869.54 - 12 \cdot (8.48)^2}{11} = 0.539$$

Como $1 - \alpha = 0.95$, $\alpha = 0.05$ y $\alpha/2 = 0.025$, buscamos en la tabla correspondiente los percentiles que se necesitan:

$$\chi^2_{11,0.025} = 21.92 \qquad \chi^2_{11,0.975} = 3.816$$

Realizando los cálculos oportunos, el IC al 95 % es:

$$(0.2708, \quad 1.5558)$$

4. Tenemos que resolver el contraste unilateral:

$$\begin{cases} H_0 & : & \mu & = & 5.9 \\ H_1 & : & \mu & > & 5.9 \end{cases}$$

Si H_0 es cierta, la medida de discrepancia será, con $\mu_0 = 5.9$:

$$d = \frac{\bar{X} - \mu_0}{\frac{s}{\sqrt{n}}} \sim t_{n-1}$$

De la muestra de 40 datos nos dan

$$\sum_{i=1}^{40} x_i = 253.5 \qquad \sum_{i=1}^{40} x_i^2 = 1628.77$$

con ellos, calculamos la media muestral y la cuasivarianza s^2:

$$\bar{x} = \frac{\sum_{i=1}^{40} x_i}{n} = \frac{253.5}{40} = 6.3375 \qquad s^2 = \frac{\sum x_i^2 - n\bar{x}^2}{n-1} = \frac{1628.77 - (40) \times (6.3375)^2}{39} = 0.569$$

Calculamos $s = 0.816$. Sustituyendo en d, obtenemos $\hat{d} = 3.606$.

Como nos dan $\alpha = 0.01$ y el contraste es unilateral por la derecha, buscamos el percentil $t_{39,0.01} = 2.426$. Como el valor de $\hat{d}$ observado es mayor que 2.426, **rechazamos la hipótesis nula para ese nivel de significación** y, por tanto, podemos suponer que la satisfacción de los clientes con la pasta de dientes es mayor ahora.

Calculamos el p-valor. En este caso es el área bajo la curva de una ditribución t_{39} a la derecha del punto $\hat{d} = 3.606$.

$$p = P(d \geq \hat{d} | H_0 \ cierta) = P(t_{39} \geq 3.606)$$

Mirando en la tabla de la t-Student vemos que esta probabilidad es menor que 0.0005 y por tanto se sitúa en la región de RECHAZO DE H_0 (es inferior a 0.01).

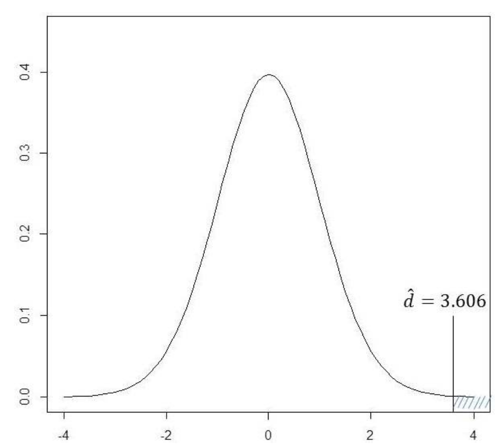

Figura 1: El p-valor es el área rayada